



Coffee Lectures

# Reusing data from the Research Collection for research, publication analysis and more

Barbara Hirschmann

1 December 2021, Webinar via Zoom

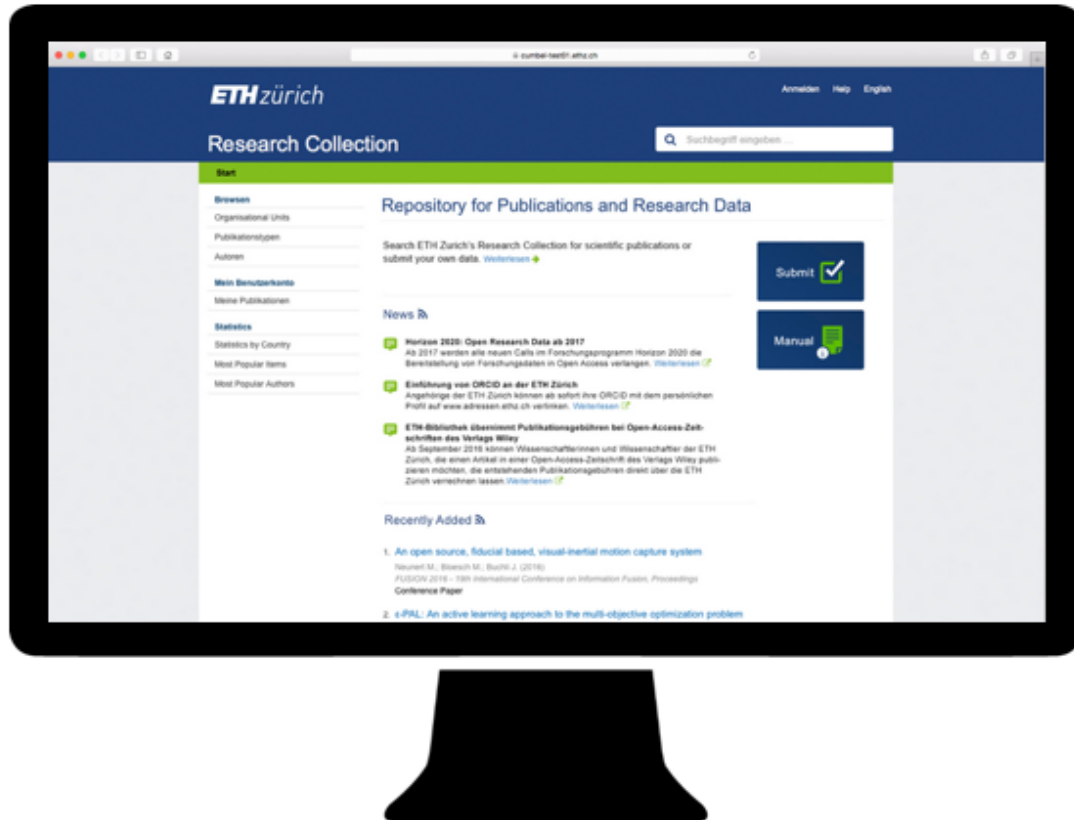
## Goals:

- understand under which conditions metadata and research data can be reused
- discover different methods to retrieve bibliographic data and usage data from the Research Collection

# Table of contents/Agenda

1. Reuse of content (publications and research data)
2. Reuse of abstracts
3. Reuse of metadata
4. Downloading metadata via the user interface
5. Retrieving metadata via an API
6. Accessing usage statistics

# Research Collection



- Publication directory / university bibliography



- Open access repository



- Research data repository

[www.research-collection.ethz.ch](http://www.research-collection.ethz.ch)

# Access modes



	Open access	Embargoed	ETHZ users	Selected users	Closed access
Publications	✓	✓			(✓)
Research data	✓	✓	✓	✓	✓

# Terms of use

**ETH zürich** Login Help Deutsch

## Research Collection

Home → Terms of Use

**Browse**

- Organisational Units
- Publication Types
- Authors

**Publish**

- New Submission

**Statistics**

- Downloads by Country
- Most Popular Items
- Most Popular Authors

### Terms of use

#### Introduction

The Research Collection is an ETH-Library service and ETH Zurich's publication platform. It offers the following features:

- **Publication database**  
The Research Collection is a comprehensive database that registers the publications produced at ETH Zurich. It is the source for publication lists
  - in ETH Zurich's academic reporting ([Annual Academic Achievements](#))
  - on ETH-Zurich websites maintained using the Content Management System [AEM](#)
- **Open access repository**  
The Research Collection is the instrument for implementing [ETH Zurich's open access policy](#). It enables members of ETH Zurich to publish academic full texts outside traditional publishing.
- **Research data archive**  
The Research Collection acts as a publication platform and digital archive for all kinds of research output. Members of ETH Zurich can
  - Publish research data as an independent publication
  - Publish research data as supplementary material for an article, dissertation or other text
  - Share research data with colleagues or a research group
  - Deposit research data for archiving purposes

#### 1. Contents

- 1.1 The Research Collection contains publications and other research output by members of staff and selected works by students from ETH Zurich (see [List publication types](#)).
- 1.2 New references are imported from bibliographic databases (currently Web of Science and Scopus) once a day. Moreover, members of staff at ETH Zurich can enter new research output independently at any time.
- 1.3 Research output that members of staff published before they joined ETH Zurich is marked accordingly in the metadata.
- 1.4 A digital object in the Research Collection might be an original work or a self-archived copy of an already published work.

#### 2. Access and re-use

- 2.1 Anyone can access bibliographic data (metadata) in the Research Collection.
- 2.2 Publications and research data can either be accessed via a digital object available in the Research Collection or via a link on an external publisher's website/repository.

# Reuse of content (publications and research data)

ETH zürich Login Help Deutsch

Research Collection Search

Home → Working Papers/Reports → Working Paper → View Item

**Browse**

- Organisational Units
- Publication Types
- Authors

**Publish**

- New Submission

**Statistics**

- Downloads by Country
- Most Popular Items
- Most Popular Authors

## The Global Economic Barometers: Composite indicators for the world economy

Download  
Full text (PDF, 3.688Mb) ↓

**Rights / license**  
In Copyright - Non-Commercial Use Permitted ↗

**Abstract**  
This paper presents a coincident and a leading composite monthly indicator for the world business cycle – the Global Economic Barometers. Both target the world's output growth rate cycle. The calculation of these indicators comprises two main stages. The first consists of a variable selection proce [Show more](#) +

**Permanent link**  
<https://doi.org/10.3929/ethz-b-000401651>

**Publication status**  
published

**Journal / series**  
KOF Working Papers

**Author**  
Abberger, Klaus   
Graff, Michael   
Campelo, Aloisio Jr.  
[Show all](#) +

**Date**  
2020-02

**Open access**

2.6 If no special end user licence is stipulated, the users may download and save the objects on offer free of charge for their own private use and for non-commercial purposes. Other uses are only possible with permission from the copyright holders.

KOF SWISS Economic Institute, ETH Zürich

# Reuse of content (publications and research data)

ETH zürich

Research Collection

Home → Research Data → Dataset → View Item

## Training Set for the Prediction of Partial Charges

Download

- ChEMBL training molecules (epsilon = 78) (Unknown, 31.22Mb) ↓
- ZINC training molecules (epsilon = 78) (Unknown, 66.00Mb) ↓
- ChEMBL training molecules (epsilon = 4) (Unknown, 31.52Mb) ↓
- ZINC training molecules (epsilon = 4) (Unknown, 67.41Mb) ↓

**Rights / license**  
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International

**Abstract**  
Parametrization of small organic molecules for classical molecular dynamics simulations is not trivial. The vastness of the chemical space makes approaches using building blocks challenging. The most common approach is therefore an individual parametrization of each compound by deriving partial charges. [Show more](#)

**Permanent link**  
<https://doi.org/10.3929/ethz-b-000230799>

**Contributors**  
Contact person: Riniker, Sereina

**Creator**  
Riniker, Sereina  
Bleiziffer, Patrick

**Date**  
2018-01

**Type**  
Dataset

**Altmetrics**  
Tweeted by 3

**Date created**  
2017

**Subject**

CC BY-NC-ND 4.0

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

This is a human-readable summary of (and not a substitute for) the license. [Disclaimer.](#)

**You are free to:**

- Share** — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

**Under the following terms:**

- Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- NonCommercial** — You may not use the material for [commercial purposes](#).
- NoDerivatives** — If you [remix](#), [transform](#), or [build upon](#) the material, you may not distribute the modified material.

**No additional restrictions** — You may not apply legal terms or [technical measures](#) that legally restrict others from doing anything the license permits.



# Reuse of abstracts


- For abstracts, the reuse conditions mentioned in the respective item under "Rights / license" apply.
- If no such rights notice is available (usually in "Metadata only" items), the abstracts may only be used **for personal use**.

The screenshot shows the ETH Zürich Research Collection interface. The breadcrumb trail is: Home → Working Papers/Reports → Working Paper → View Item. The main title is "The Global Economic Barometers: Composite indicators for the world economy". A red box highlights the "Rights / license" section, which states: "In Copyright - Non-Commercial Use Permitted". Other visible sections include "Download" (Full text PDF, 3.688Mb), "Abstract" (This paper presents a coincident and a leading composite monthly indicator for the world business cycle...), "Author" (Abberger, Klaus; Graff, Michael; Campelo, Aloisio Jr.), "Date" (2020-02), "Type" (Working Paper), and "Publisher" (KOF Swiss Economic Institute, ETH Zurich).

The screenshot shows the ETH Zürich Research Collection interface. The breadcrumb trail is: Home → Journal Contributions → Journal Article → View Item. The main title is "On the Capacity of MIMO Optical Wireless Channels". A red box highlights the "Abstract" section, which states: "This paper studies the capacity of a general multiple-input multiple-output (MIMO) free-space optical intensity channel under a per-input-antenna peak-power constraint and a total average-power constraint over all input antennas. The focus is on the scenario with more transmit than receive antennas Show more". Other visible sections include "Publication status" (published), "External links" (https://doi.org/10.1109/TIT.2020.2979716), "Journal / series" (IEEE Transactions on Information Theory), "Volume" (66(9)), "Pages / Article No." (5660 - 5682), and "Publisher" (Institute of Electrical and Electronic Engineers).

# Reuse of metadata


- All metadata (excl. abstracts) are available for reuse under a [Public Domain Dedication \(CC0\)](#).
- They may be used, modified and redistributed without prior consent.




**CC0 1.0 Universal (CC0 1.0)  
Public Domain Dedication**

This is a human-readable summary of the [Legal Code \(read the full text\)](#). [Disclaimer](#)

### No Copyright

 The person who associated a work with this deed has **dedicated** the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See **Other Information** below.



# Downloading metadata via the user interface

The screenshot shows the 'Research Collection' user interface. At the top, there is a search bar with the text 'Suche'. Below it, a navigation bar shows 'Start → Suche'. On the left, there are menu sections for 'Browsen', 'Publizieren', and 'Statistik'. The main content area is titled 'Suche' and shows a search input field containing the query 'ethz.leitzahl:03565 and dc.type:"journal article"'. A red arrow points to this search input. Below the search bar, there is a section for 'Ergebnisse' (Results) showing 'Anzeige der Einträge 1-10 von 320'. A red arrow points to the 'Export' button, which has a dropdown menu open. The dropdown menu lists three options: 'CSV' (320), 'RIS' (320), and 'BibTeX' (320). Below these options is a link 'Treffer auswählen'. To the right of the search results, there is a 'Filtern nach' (Filter by) section with a list of authors and their counts, such as 'Müller, Ralph (270)', 'van Lenthe, G. Harry (51)', etc.

- Use [expert search](#) terms
- Download lists with less than 500 items
- For larger lists: separate into smaller result sets
- Or: contact [research-collection@library.ethz.ch](mailto:research-collection@library.ethz.ch) for lists with more than 500 items

# Retrieving metadata via APIs

## OAI-PMH interface

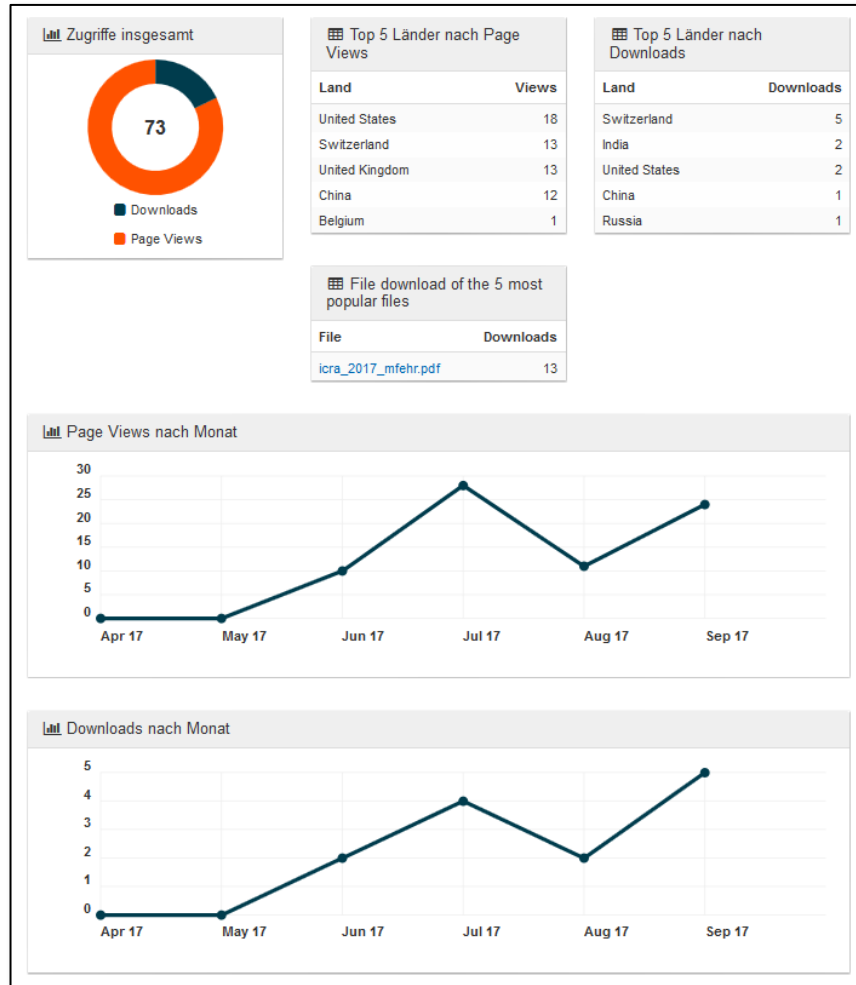
- For incremental harvesting
- Different standardised metadata formats available (Dublin Core, DataCite,...)
- Different subsets of items available (all items, only items with fulltext/object, only certain publication types)
- Public access
- Documentation available at: <https://documentation.library.ethz.ch/display/RC/OAI-PMH+interface>

## REST API

- For automated searches, publication lists (on webpages)
- ETHZ specific metadata format based on Dublin Core (XML oder JSON)
- Results can be filtered by date, author, organizational unit,...
- Access for members of ETH Zurich only
- Documentation available on request

The screenshot displays the DSpace OAI-PMH Data Provider interface. At the top, there is a navigation bar with links for 'Identify', 'Sets', 'Records', 'Identifiers', and 'Metadata Form'. Below this, a 'List of Records' section shows 'Results fetched 0 - 10 of 63514'. A specific record is highlighted with the following details: Identifier: oai:www.research-collection.ethz.ch:20.500.11850/52, Last Modified: 2021-07-30 15:41:23. Below the record information, there are sections for 'Sets' and 'Metadata'. The 'Metadata' section displays XML data for a 'Doctoral Thesis' by Gasparini, Blaž, titled 'Cirrus clouds and their geoengineering potential', issued in 2016. The metadata includes identifiers for the document, publisher (ETH Zurich), and license (In Copyright - Non-Commercial Use Permitted), along with contributor information and a detailed abstract.

# Accessing usage statistics



- Usage statistics are available on item level for individual users
- The Research Collection team can trigger an export of **downloads and page views for a list of items** specified by e.g. Leitzahl, publication date, series title,...
- Contact us at [research-collection@library.ethz.ch](mailto:research-collection@library.ethz.ch)

# Your support services at ETH Library

- Research Collection: [research-collection@library.ethz.ch](mailto:research-collection@library.ethz.ch) for
  - batch export of search results > 500 items
  - batch export of usage statistics
  - any other Research-Collection-related support requests
- Bibliometrics: [bibliometrics@library.ethz.ch](mailto:bibliometrics@library.ethz.ch) for
  - impact analysis (group-level)
  - generation of bibliometric profiles and reports
  - individual advice on the topic of bibliometrics
- Text and data mining: [eressourcen@library.ethz.ch](mailto:eressourcen@library.ethz.ch) for
  - text and data mining projects based on the media licensed by ETH Library



Thanks for your attention!

Join us for our next Coffee Lectures  
in spring 2022!

Watch: [www.library.ethz.ch/coffee-lectures](http://www.library.ethz.ch/coffee-lectures) for program details!

Barbara Hirschmann  
Head E-Publishing  
barbara.hirschmann@library.ethz.ch

ETH Zurich  
ETH Library  
Rämistrasse 101  
8092 Zurich

[www.library.ethz.ch/coffee-lectures](http://www.library.ethz.ch/coffee-lectures)